Least squares coefficients for a quadrature formula for Laplace transform inversion

Carl. P. Jeffreson and Ee-Pan Chow (*)

ABSTRACT

A linear iterative method of least squares approximation of functions by exponentials due to Miller [9] is adapted to derive a set of least squares coefficients for an approximate Laplace transform inversion formula eq. (1). An earlier assumption made by Zakian [2] - that the approximation to the Laplace transform inverse will improve provided the approximation to the Dirac delta function is improved - is shown to be not substantiated for a number of test functions.

1. INTRODUCTION

Let F(s) be the Laplace tranform of f(t).
A well-known formula for the approximate inversion of F(s) is

$$f_n(t) = \frac{1}{t} \sum_{i=1}^{n} K_i\, F\left(\frac{\alpha_i}{t}\right) \qquad 0 < t < \infty \tag{1}$$

where $f_n(t)$ is an approximation of f(t).
A number of methods can be invoked to determine values of the coefficients $\{\underset{\sim}{\alpha}, \underset{\sim}{K}\}$. Salzer [3] computed these coefficients so that $f_n(t) = f(t)$ whenever $F(s) = s^{-k}$, k = 1, 2, ..., 2n. Then eq. (1) is in fact a Gaussian quadrature of Bromwich's integral. Salzer's formula is generalized by Krylov and Skoblya [17]. Vlach [5, 6] and Zakian and Edwards [2, 4] have obtained a similar set of 'Gaussian' coefficients, using an approach based on Padé approximation. Zakian [1] has shown that the set of coefficients $\{\underset{\sim}{\alpha}, \underset{\sim}{K}\}$ has to be determined so that

$$\delta_n(\tau-1) = \sum_{i=1}^{n} K_i \exp(-\alpha_i\tau) \quad (\alpha_i\text{'s distinct}) \tag{2}$$

is a good approximation of the Dirac delta function $\delta(\tau-1)$ for $0 \leqslant \tau \leqslant \infty$. Several approximation criteria can be used. Zakian and Gannon [7] derived a set of quasi least-squares coefficients $\{\underset{\sim}{K}\}$ for a chosen set of real valued $\{\underset{\sim}{\alpha}\}$, which however, when applied to eq. (1) gave worse results than the complex Gaussian coefficients for a given n [8].
In this paper, we present a set of "true" least-squares (LS) coefficients $\{\underset{\sim}{\alpha}, \underset{\sim}{K}\}$. The method of LS approximation of functions by exponentials proposed by Miller [9] is adapted to derive a set of coefficients $\{\underset{\sim}{\alpha}, \underset{\sim}{K}\}$ so that $\delta_n(\tau-1)$ approximates $\delta(\tau-1)$ in a least-squares sense. The problem is constrained so that eq. (1) gives $f_n(\infty) = f(\infty)$. The validity of Zakian's assumption - that $f_n(t)$ will be a better approximation to f(t) provided the approximation $\delta_n(\tau-1)$ to $\delta(\tau-1)$ is improved - is also tested by comparing $f_n(t)$ obtained using true LS coefficients (complex) with those obtained using Gaussian coefficients, for a number of commonly-used transform pairs having known analytical inverses.

2. CONSTRAINED LEAST-SQUARES APPROXIMATION OF A SQUARE PULSE FUNCTION

To derive a set of LS coefficients $\{\underset{\sim}{\alpha}, \underset{\sim}{K}\}$ we seek to minimise

$$E(\underset{\sim}{\alpha}, \underset{\sim}{A}) = \int_0^\infty \{f(\tau) - \sum_{i=1}^{n} A_i\, e^{-\alpha_i\tau}\}^2\, d\tau. \tag{4}$$

Since the integral does not exist for $f(\tau) = \delta(\tau-1)$, we consider the square pulse function $f(\tau) = m(\tau-1)$ (as Zakian and Gannon did [7]), where

$$m(\tau-1) = 1 - \int_0^\tau \delta(\theta-1)\, d\theta = \begin{cases} 1, & 0 \leqslant \tau \leqslant 1 \\ 0, & \tau \geqslant 1 \end{cases} \tag{5}$$

which gives a finite integral in eq. (4). An implicit assumption here is that improving the square pulse function approximation will also result in improving the delta function in some way.
Similarly,

$$m_n(\tau-1) = 1 - \int_0^\tau \delta_n(\theta-1)\, d\theta = \sum_{i=1}^{n} \frac{K_i}{\alpha_i} e^{-\alpha_i\tau} \tag{6}$$

provided,

$$\sum_{i=1}^{n} \frac{K_i}{\alpha_i} = \sum_{i=1}^{n} A_i = 1. \tag{7}$$

(*) C. P. Jeffreson, Ee-Pan Chow, Dept. of Chemical Engineering, University of Adelaide, North Terrace, Adelaide, South Australia.

Condition (7) ensures that the initial and final values of the function $f_n(t)$ are equal to those of the function f(t) [8]. Minimisation of $E(\underset{\sim}{a}, \underset{\sim}{A})$ for the square pulse function $m(\tau\text{-}1)$ subject to the linear constraint (7) can be accomplished using the Lagrange multiplier optimisation technique [10]. Following Miller's procedure [9], we derived sets of LS coefficients $\{\underset{\sim}{a}, \underset{\sim}{A}\}$ for the square pulse function $m(\tau\text{-}1)$ (see Appendix A) for up to n = 15 terms using double precision arithmetic. It was found that Miller's algorithm converges very slowly. However the speed of convergence can be improved considerably by using the arithmetic means of the last two iteration values of each a_i after every two or three iterations using Miller's iterative scheme. Values of LS coefficients $\{\underset{\sim}{a}, \underset{\sim}{A}\}$ for the unconstrained case obtained using this refinement to Miller's algorithm match those published by Miller [9] and Longman [11]. Values of the constrained LS coefficients $\{\underset{\sim}{a}, \underset{\sim}{K}\}$ for n = 5, 10 and 15 (with 10 s.f. convergence) are shown in Table 1. Stroud and Secrest [12] tabulate Gaussian coefficients for several values of n. It is worth noting the small moduli of the K_i's of the LS coefficients. This gives the LS coefficients an advantage over the Gaussian coefficients which have large moduli for the K_i's, rendering them vulnerable to numerical instability and the loss of arithmetic precision, especially for large values of n.

Figure 1 compares the integral square errors (ISE's) of the approximations of the square pulse function using LS-constrained and unconstrained - and Gaussian coefficients for n up to 15 terms. (The rational function used in the Padé approximation to determine the Gaussian coefficients has a numerator polynomial with degree one less than that of the denominator polynomial). It can be seen that the constrained LS approximations give slightly higher ISE's than the unconstrained case, but considerably better ISE's than the approximations using Gaussian coefficients. The constrained LS approximation gives a 38 % improvement in ISE over the Gaussian results for n = 15. If Zakian's hypothesis is correct, then these results suggest that, for a given n, using the constrained LS coefficients in eq. (1) will produce a better approximation $f_n(t)$ to f(t) than using the Gaussian coefficients.

3. LAPLACE INVERSION USING COEFFICIENTS

To test Zakian's hypothesis, eq. (1), using both constrained LS coefficients and Gaussian coefficients for n = 15, was invoked to invert five test functions. Table 2 lists the test functions and their analytical inverses, together with a comparison of the ISE's between the approximations and exact functions. The actual errors of both approximations - using

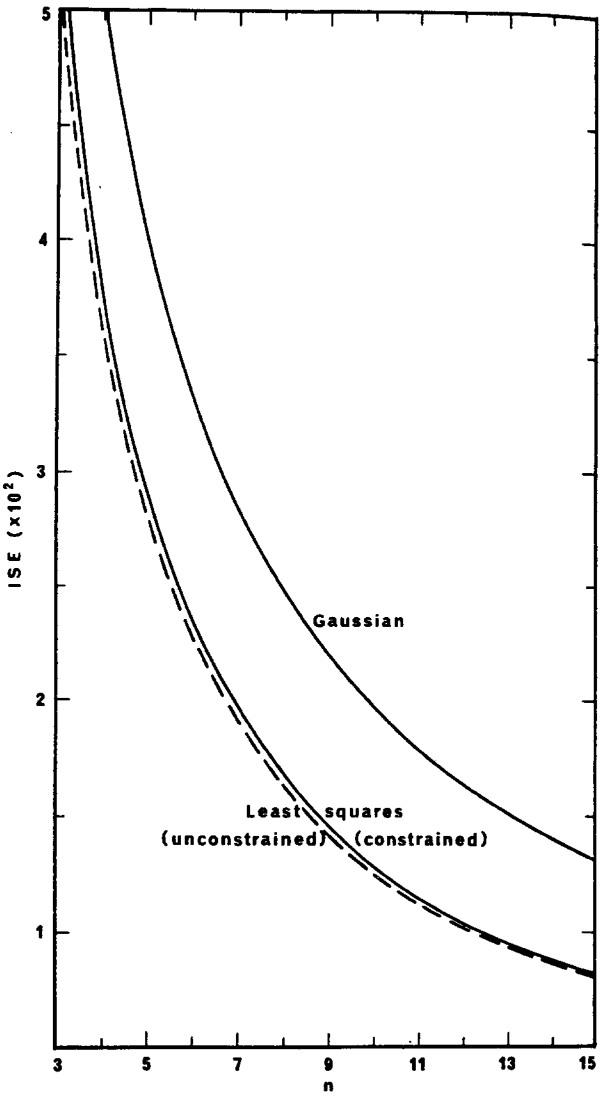

Fig. 1. Plot of integral square errors (ISE) against n for a square pulse function.

LS and Gaussian coefficients for n = 15 - are plotted against time for each of the five test functions in Fig. 2. The plots indicate that, in each case, the actual errors and ISE's of the approximation using the Gaussian coefficients are lower than those obtained using LS coefficients. The prima facie reason is perhaps that a better approximation of a square pulse function does not necessarily result in a better approximation of the Dirac delta function. However, following Longman's definition [11], a comparison of the "partial" ISE's of the Dirac function approximation still suggests that the LS approximation is a better approximation than the Gaussian approximation (see Table 3). We have now shown that using a set of coefficients derived through a more accurate representation of the square pulse function (or the delta function) need not necessarily result in a more accurate approximation to f(t). In fact, for the examples used, the Gaussian coefficients yield far more accurate approximations than the LS coefficients.

Table 1. Least squares coefficients for equation (1)

REAL(ALPHA)	IMAG(ALPHA)	REAL(K)	IMAG(K)	(*)
N = 5				
1.313078739E+00	-9.180380688E+00	-2.675601562E+00	6.122379137E-02	(2)
2.243145857E+00	-4.382349912E+00	-1.033032040E+00	7.130785840E+00	(2)
2.807495678E+00	0.	1.085010133E+01	0.	(1)
N = 10				
1.230093058E+00	-2.352672861E+01	4.860502512E-01	2.422640698E+00	(2)
1.948441344E+00	-1.838606040E+01	5.552772067E+00	2.015535066E+00	(2)
2.486321472E+00	-1.294069981E+01	9.328911705E+00	-4.782670037E+00	(2)
3.049851019E+00	-7.477695691E+00	4.688193150E+00	-1.756397771E+01	(2)
3.662673996E+00	-2.340402588E+00	-2.290775011E+01	-1.934430556E+01	(2)
N = 15				
1.201900131E+00	-3.839128154E+01	1.454512148E+00	-1.896165496E+00	(2)
1.862570769E+00	-3.320749609E+01	-1.715333565E+00	-5.172539073E+00	(2)
2.292993575E+00	-2.766209815E+01	-7.541523912E+00	-4.523961971E+00	(2)
2.661840425E+00	-2.198372929E+01	-1.295649648E+01	7.506458220E-01	(2)
3.035577722E+00	-1.625582381E+01	-1.547416249E+01	1.099375660E+01	(2)
3.470465540E+00	-1.055035834E+01	-1.018991823E+01	2.717989379E+01	(2)
4.014204990E+00	-5.028754130E+00	1.938392606E+01	4.294784889E+01	(2)
4.382910986E+00	0.	6.166590165E+01	0.	(1)

* VARIABLES THAT OCCUR IN COMPLEX CONJUGATE PAIRS ARE DESIGNATED WITH (2)S.

Table 2. List of test functions

Case	F(s)	f(t)	ISE* (least squares)	ISE* (Gaussian)	T≠	Reference
1	$1/(s+1)$	$\exp(-t)$	3.25E-05	1.02E-24	7.0	[2]
2	$\exp(-\sqrt{s}/2)$	$\frac{1}{4\sqrt{\pi t^3}} \exp(\frac{-1}{16t})$	1.73E-02	2.51E-03	1.50	[13]
3	$1/(s^2+s+1)$	$\frac{2}{\sqrt{3}} \exp(\frac{-t}{2}) \sin(\frac{\sqrt{3}}{2}t)$	3.22E-04	1.03E-18	15.0	[14]
4.	$1/\{s(\sqrt{s}+1)$	$1 - \exp(t)\, \mathrm{erfc}(\sqrt{t})$	3.70E-04	9.57E-06	10.0	[14]
5	$\exp\{5(1-\sqrt{1+2s/5}\}/s$	$\frac{1}{2}\{\mathrm{erfc}[\frac{1}{2}\sqrt{\frac{10}{t}}\,(1-t)]$ $+ e^{10}.\mathrm{erfc}[\frac{1}{2}\sqrt{\frac{10}{t}}\,(1+t)]\}$	3.25E-07	1.03E-08	4.0	[8]
	(step response of axial dispersion model with simplified boundary conditions for Pe=10)					

$$* \quad \mathrm{ISE} = \sum_{t=0}^{t=T} \Delta t\, \{f_n(t)-f(t)\}^2 \text{ for } n = 15.$$

≠ The range of summation is $0 \leqslant t \leqslant T$.
Although eq.(1) cannot be used for t=0, we can still compute $f_n(t=0)$ using the initial value theorem:

$$\lim_{t=0} f_n(t) = \sum_{i=1}^{n} \frac{K_i}{\alpha_i} \lim_{s_i\to\infty} s_i\, F(s_i) = \lim_{t\to 0} f(t) \ ,$$

provided $\sum_{i=1}^{n} \frac{K_i}{\alpha_i} = 1.$

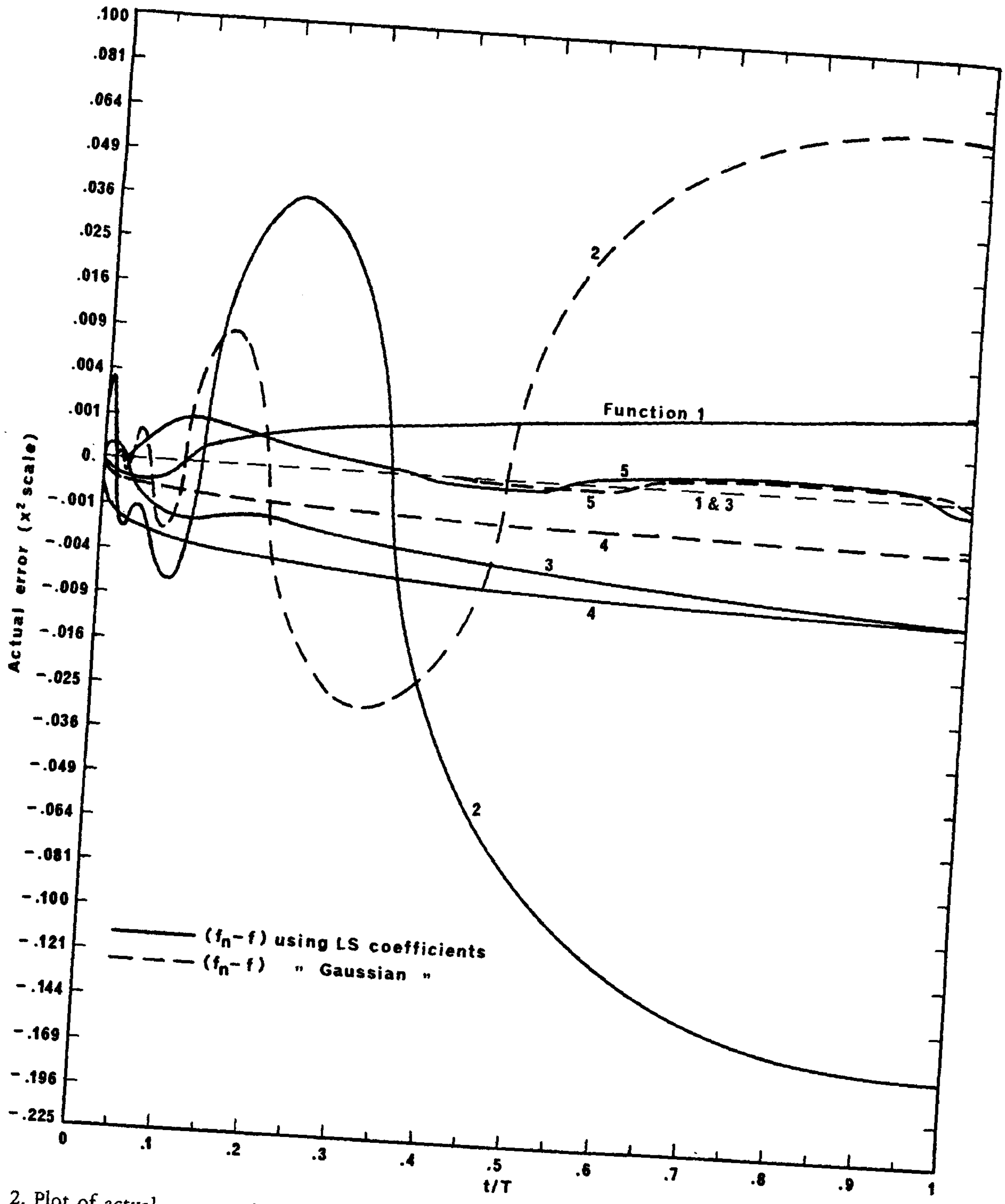

Fig. 2. Plot of actual errors against t/T for n = 15. (The test functions and T's used are shown in Table 2)

4. CONCLUSIONS

Miller's method of least squares approximation of functions by exponentials can be adapted to derive a set of "true" least squares coefficients for the approximate Laplace inversion formula eq. (1). These LS coefficients have the advantage of small moduli for the K_i's, thus affording better numerical stability to eq. (1) than the Gaussian coefficients. Although the square pulse function (or the delta function) can be more accurately approximated using least square approximants than Gaussian approximants, their corresponding coefficients, when used in eq. (1), do not necessarily produce more accurate approximations to f(t). Our results seem to cast doubt on Zakian's assertion that "the more nearly $\delta_n(\tau-1)$ behaves like the Dirac delta function $\delta(\tau-1)$, the more accurately $f_n(t)$ approximates f(t)" [2]. Although Zakian's assumption is not strictly valid, eq. (1) is

Table 3. A comparison of approximations of $\delta(\tau-1)$

n	ISE(δ)* (least squares)	ISE(δ)* (Gaussian)
5	- 3.19	-1.62
10	- 7.65	-3.52
15	-12.31	-5.47

$$* \ \mathrm{ISE}(\delta) = \int_0^\infty \{\delta(\tau-1) - \sum_{i=1}^{n} K_i e^{-\alpha_i \tau}\}^2 d\tau - \int_0^\infty \{\delta(\tau-1)\}^2 d\tau$$

$$= -2 \sum_{i=1}^{n} K_i e^{-\alpha_i} + \sum_{i=1}^{n} \sum_{k=1}^{n} \frac{K_i K_k}{(\alpha_i+\alpha_k)}$$

still very useful as a numerical method for Laplace inversion, and fairly accurate results can be obtained using the Gaussian coefficients.

5. ACKNOWLEDGEMENT

We thank the referee for his helpful comments.

REFERENCES

1. ZAKIAN, V. : "Numerical inversion of Laplace transform", Elec. Lett., *5* (1969), pp. 120-121.
2. ZAKIAN, V. : "Optimisation of numerical inversion of Laplace transforms", Elec. Lett., *6* (1970), pp. 677-679.
3. SALZER, H. E. : "Orthogonal polynomials arising in the numerical evaluation of inverse Laplace transforms", MTAC, *9* (1955), pp. 164-177.
4. ZAKIAN V. and EDWARDS M. J. : "Tabulation of constants for full grade I_{MN} approximants", Control System Centre Report No. 312, Univ. of Manchester Institute of Science and Technology, (June 1976).
5. VLACH, J. : "Numerical method for transient responses of linear networks with lumped, distributed or mixed parameters", J. Franklin Inst., *288* (1969), pp. 99-113.
6. SINGHAL, K. and VLACH, J. : "Computation of time domain response by numerical inversion of the Laplace transform", J. Franklin Inst., *299* (1975), pp. 109-126.
7. ZAKIAN, V. and GANNON, D. R. : "Least-squares optimisation of numerical inversion of Laplace transforms", Elec. Lett., 7 (1971), pp. 70-71.
8. CHOW, E. P. and JEFFRESON, C. P. : "Application of numerical Laplace transform inversion techniques to the analysis of the dynamics of short parallel plate packed beds", Chem. Eng. in Austr., *Ch. E. 2* (1977), pp. 11-15.
9. MILLER, G. : "Least squares approximation of functions using exponentials", Ph. D. thesis, John Hopkins Univ., (1969).
10. CURTIS, P. C., Jr., *Multivariate calculus with Linear Algebra*, Wiley, N. Y. (1972), p. 289.
11. LONGMAN, I. M. : "Application of best rational function approximation for Laplace transform inversion", Journal of Computational and Applied Mathematics, *1* (1975), pp. 17-23.
12. STROUD, A. H. and SECREST, D. : *Gaussian Quadrature Formulas*, Prentice-Hall, N. J. (1966).
13. MILLER, M. K. and GUY W. T. : "Numerical inversion of the Laplace transform by use of Jacobi polynomials", SIAM J. Numer. Anal., *3* (1966), p. 624.
14. WEEKS, W. T. : "Numerical inversion of Laplace transform using Laguerre functions", J. ACM, *13* (1966), pp. 419-426.
15. GASTINEL, N. : "Inversion d'une matrice généralisant la matrice de Hilbert", Chiffres, *3* (1960), pp. 149-152.
16. TOU, J. L. : "Determination of the inverse Vandermonde matrix", IEEE Trans. on Automation Control, AC-9 (1964), p. 314.
17. KRYLOV, V. I. and SKOBLYA, N. S. : "On the numerical inversion of the Laplace transform", Inzh-Fiz. Zh., *4* (1961), pp. 85-101.

APPENDIX
MILLER'S LINEAR ITERATIVE SCHEME APPLIED TO A CONSTRAINED PROBLEM

We seek to minimise

$$E(\underset{\sim}{a}, \underset{\sim}{A}) = \int_0^\infty \{m(\tau-1) - \sum_{i=1}^{n} A_i e^{-a_i\tau}\}^2 d\tau \qquad (A1)$$

subject to the linear constraint

$$g(\underset{\sim}{A}) = \sum_{i=1}^{n} A_i = 1. \qquad (A2)$$

Using the Lagrange multiplier method [10], we require

$$\underset{\sim}{\nabla} E(\underset{\sim}{a}, \underset{\sim}{A}) = \lambda \nabla g(\underset{\sim}{a}, \underset{\sim}{A}) \qquad (A3)$$

where λ is a scalar (Lagrange multiplier).
Eq. (A3) is satisfied if and only if

$$\left.\begin{array}{l} \sum_{k=1}^{n} 2A_k \int_0^\infty e^{-(a_i+a_k)\tau} d\tau - 2\int_0^\infty f(\tau)\, e^{-a_i\tau} d\tau = \lambda \\ \text{and } -\sum_{k=1}^{n} 2A_i A_k \int_0^\infty \tau e^{-(a_i+a_k)\tau} d\tau + 2A_i \int_0^\infty \tau\, f(\tau)\, e^{-a_i\tau} d\tau = 0 \end{array}\right\} \text{for } i=1,. \quad (A4)$$

Following Miller's method [9], we write eq. (A4) in the frequency domain,

$$\left.\begin{array}{l} \sum_{k=1}^{n} \frac{A_k}{(a_i+a_k)} = \frac{\lambda}{2} + F(a_i) \\ \text{and } -\sum_{k=1}^{n} \frac{A_k}{(a_i+a_k)^2} = F'(a_i) \end{array}\right\} \text{for } i=1,\ldots,n. \qquad (A5)$$

Writing $F_n(s) = \sum_{k=1}^{n} \frac{A_k}{(s+s_k)}$ and $F_n'(s) = -\sum_{k=1}^{n} \frac{A_k}{(s+s_k)^2}$,

eq. (A5) simplifies to

$$F_n(a_i) = \frac{\lambda}{2} + F(a_i) \quad \text{and} \quad F_n'(a_i) = F'(a_i) \quad \Bigg\} \text{ for } i = 1,\ldots,n. \tag{A6}$$

If we write

$$F_n(s) = \frac{a_1 + a_2 s + \ldots + a_n s^{n-1}}{b_1 + b_2 s + \ldots + b_n s^{n-1} + s^n} = \frac{N(s)}{D(s)} \tag{A7}$$

$$\left.\begin{array}{l} \text{then } (\frac{\lambda}{2} + F)\, D = N \\ \text{and } F'D + (\frac{\lambda}{2} + F)\, D' = N' \end{array}\right\} \text{for } s = a_i, i=1,\ldots,n \tag{A8}$$

where $-a_i$, $i=1,\ldots,n$, are the roots of the n^{th} degree polynomial $D(s)$. Substituting eq. (A7) into eq. (A8) gives a set of $2n$ equations, which may be written in matrix form as

$$\left.\begin{array}{l} [V]\, \underset{\sim}{a} + [G]\, \underset{\sim}{b} = \underset{\sim}{X} \\ \text{and } [W]\, \underset{\sim}{a} + [P]\, \underset{\sim}{b} = \underset{\sim}{Y} \end{array}\right] \tag{A9}$$

where $v_{ik} = (a_i)^{k-1}$, $g_{ik} = -(a_i)^{k-1}\{\frac{\lambda}{2} + F(a_i)\}$,

$w_{ik} = (k-1)(a_i)^{k-2}$, $p_{ik} = -(a_i)^{k-1}F'(a_i)$
$- (k-1)\,(a_i)^{k-2}\{\frac{\lambda}{2} + F(a_i)\}$,

$x_i = (a_i)^n\,\{\frac{\lambda}{2} + F(a_i)\}$, $y_i = (a_i)^n\,F'(a_i)$
$+ n(a_i)^{n-1}\,\{\frac{\lambda}{2} + F(a_i)\}$,

$\underset{\sim}{a} = \{a, a_2 \ldots a_n\}^T$ and $\underset{\sim}{b} = \{b_1\ b_2 \ldots b_n\}^T$,
for $i = k = 1, \ldots, n$.

Solving eqns. (A9) simultaneously gives

$$\underset{\sim}{b} = [[P] - [W][V]^{-1}[G]]^{-1}\ \{\underset{\sim}{Y} - [W][V]^{-1}\underset{\sim}{X}\}. \tag{A10}$$

The Lagrange multiplier λ may be computed from eq. (A5a), which gives, in matrix form,

$$\underset{\sim}{A} = [H]^{-1}\ \{\frac{\lambda}{2}\underset{\sim}{I} + \underset{\sim}{F}(a)\}, \tag{A11}$$

where $[H]$ is the Hilbert matrix with elements $h_{ik} = 1/(a_i + a_k)$ and has an explicit inverse [14], and $\underset{\sim}{I}$ is the unity matrix.
To satisfy the constraint (A2), we must have

$$\sum_{i=1}^{n}\sum_{k=1}^{n} h_{ik}^{-1}\ \{\frac{\lambda}{2} + F(a_k)\} = 1, \tag{A12}$$

from which we can solve for $\lambda/2$ explicitly.
We now have a linear iterative scheme, from which we can compute $\underset{\sim}{b}$ from eq. (A10) at each iteration using the set of a_i's - the negative values of the roots of $D(s)$ - of the previous iteration. Finally the coefficients $\underset{\sim}{A}$ are obtained from eq. (A11) using the current values of $\underset{\sim}{a}$. Note that $[V]$ is the Vandermonde matrix, which has an analytical inverse given by Tou [16]. For the square pulse function $m(\tau-1)$,

$$F(s) = \frac{1-e^{-s}}{s}, \quad F'(s) = \frac{1}{s}\,\{e^{-s} - F(s)\}, \tag{A13}$$

and the integral square error (ISE) for its approximation can be computed from the equation

$$ISE = 1 - 2\sum_{i=1}^{n}\frac{A_i}{a_i}\,(1-e^{-a_i}) + \sum_{i=1}^{n}\sum_{k=1}^{n}\frac{A_i A_k}{(a_i+a_k)}. \tag{A14}$$